Dense Matching Methods for 3D Scene Reconstruction from Wide Baseline Images

Zoltán Megyesi

PhD Theses

Supervisor: Prof. Dmitry Chetverikov

Eötvös Loránd University
PhD Program in Informatics
Program Director: Prof. János Demetrovics

Computer and Automation Research Institute
Hungarian Academy of Sciences

Faculty of Mechanical Engineering and Automation
Kecskemét College

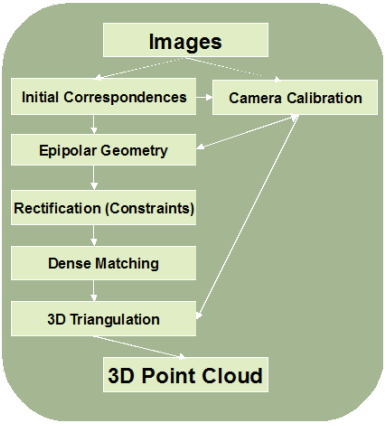
Budapest 2009

1 Introduction

The topic of this dissertation is *3D scene reconstruction*, which is a quickly evolving field of Computer Vision. The reconstruction is the first step of 3D acquisition related processes, its purpose is to measure 3D scenes and objects to create 3D data for the use of 3D applications. The data that is measured is a 3D point cloud representing the scene, therefore reconstruction is often called passive visual 3D scanning.

What makes reconstruction special among 3D scanners is that it creates 3D data purely from visual signs. This feature ensures a wide applicability, as visual information is relatively easy and cheap to come by and it can provide abundant 3D information. The visual source makes 3D reconstruction especially useful for tasks where realistic 3D data is used for visualization purposes, ranging from archeology to entertainment.

The input for the reconstruction is images from different viewpoints. The 3D information is acquired from those pixels of the different images that are projections of the same 3D points. In the first step we need to estimate some initial geometry of the images and the cameras. This geometry can be used to perform *dense matching*, the task that tries to match every visible corresponding pixel across the images. With the knowledge of accurate camera information, the dense correspondence set can be turned into a 3D point cloud using 3D triangulation. The summary of the process can be seen below.



In the dissertation, we discuss the whole 3D scene reconstruction process, but focus on image processing and dense matching. We emphasize the special case when the *baseline* of the cameras is *wide*. This case has the advantage of increased accuracy, but makes matching more challenging due to the large distortion between the views.

2 Outline of the dissertation

After the introduction in *Chapter 1*, we provide an overview of the reconstruction problem including scene types, setups, requirements and finally the summary of a general multi-step solution in *Chapter 2*. We discuss the subtask of dense matching in more details in *Chapter 3*. We cover similarity functions, constraints for search space limitation and overview different frameworks. We also provide a pseudo code for region growing based dense matching. In *Chapter 4* we discuss the wide-baseline setup and introduce Dense Affine Matching as a dense matching solution applicable on wide baseline images. In *Chapter 5* a new method is provided to calculate surface normals from the distortion of image patches, and the Normal Aided Matching is introduced. The description of an implemented software system is provided in *Chapter 6*, which is followed by the collection of test data sets and results of different experiments in *Chapter 7*. A new quantitative evaluation method for dense matching is introduced in *Chapter 8* and the evaluation of the contributed dense matching algorithms is provided. The conclusion and future work can be found in *Chapter 9* along with the contributions of the study and the publications of the author.

3 Summary of contributions of the dissertation

In this dissertation I present a new way to discuss the properties of the dense matching algorithms, based on the way they constrain the search in the matching. I developed a region growing framework for dense matching algorithms that can apply several constraints and can adopt different algorithms, including the advanced methods discussed in this dissertation. Important constraints and the framework are discussed in *Chapter 3* and they have been presented in [1, 3, 6].

During my experiments with wide baseline images, I observed that the main source of errors come from the typically large distortion between views. I designed a matching function that can compensate the distortion by affine transformations, and used it to formulate novel constraints and a new region growing based dense matching algorithm (Dense Affine Matching, DAM). I tested the algorithm on both real world and synthetic data sets and evaluated the results. When compared to an efficient classical method, the new algorithm proved superior in terms of accuracy, density and reliability. The DAM method is discussed in *Chapter 4*, with its evaluation in *Chapter 8*. The method was presented in [3, 6, 7].

I examined the connection between the distortion and the surface orientation, and found a way to calculate normals from affine transformations using camera calibration data. I used the surface normals to form geometrically more correct constraints and applied them in a novel dense matching algorithm (Normal Aided Matching). With this matching method I found a new way to use and exploit calibration data in the dense matching. I also tested this algorithm on both

real world and synthetic data sets and evaluated the results. When compared to both DAM and the same efficient classical method, the new algorithm again proved superior in terms of both accuracy and reliability. The method is discussed in *Chapter 5*, with its evaluation in *Chapter 8*. The method was presented in [8, 9] and was used in [11, 13].

The quantitative evaluation of dense matching algorithms is not solved yet. It is hard to find testbeds and evaluation tools that are also good for wide baseline methods. I designed a new evaluation scheme that is capable of comparing dense matching methods (including wide baseline ones) in a quantitative manner. The evaluation takes a groundtruth data to separate the reconstructed point cloud to inliers and outliers using LMedS. The important figures are the accuracy of the inliers, and the number of the outliers. The evaluation method is discussed in *Chapter 8*. The method was first presented in [7] but was also used for evaluation in [8, 9].

I collected and created data for an evaluation testbed. Each data set contained a groundtruth model, snapshots taken from different views, and accurate calibration data. Initially only semi-synthetic data (by photo-realistic texturing of Janko et al.) was used, but there are possibilities for inserting real world data in the testbed. These data sets were used to evaluate the introduced dense matching methods and along with real life data, they are presented in *Chapter 7*. These data sets and results appear throughout the related publications of the author.

I collected and implemented all missing steps to create a complete

reconstruction software system (SceneRec). I designed the software to be modular, for better flexibility and to fit its purpose in research and education better. The modules of the implementation are based on the multi-step process discussed in the study. During the implementation different tools and programming languages were used. The system has a graphical interface and a script library to access the advanced functions. The different modules communicate with a clear and understandable interface of images and ASCII metadata files. The SceneRec system was used to test the dense matching methods and to perform 3D reconstruction. The module implementations are discussed in *Chapter 6* and the results can be seen in *Chapter 7*. The software system was used to provide results in all of the authors related publications.

4 New scientific results

The main contributions of the dissertation are summarized in the following three theses.

Thesis 1 Modular Software System for Reconstruction and Evaluation Testbed

I have developed a modular software system for all components of the reconstruction process. I utilize a novel way to describe properties of dense matching through constraints. The system also includes a testbed and a method for quantitative evaluation of dense matching algorithms.

1.1 I have created a software system (ScenerRec) that contains all necessary steps of the reconstruction process. The system is modular to have flexibly replaceable components and for providing solutions for different acquisition setups. The system has a simple graphical interface, uses multiple programming languages for the components and has a transparent interface of text files and images. The system can output 3D data in a standard format for visualization and further use.

1.2 I introduced a general description for the key distinguishing feature of different dense matching algorithms, namely the way the methods reduce search space. This description was expressed in the form of constraints, and was used to describe the properties of the matching algorithm.

1.3 I have created a region growing based dense matching framework that is capable of utilizing different matching methods, including some classical ones. The framework was shown to be capable of incorporating several important constraints.

1.4 I have created a realistic semi-synthetic dataset for testing dense matching algorithms containing groundtruth, virtual images and accurate calibration data. The testbed has been used to test different algorithms.

1.5 I have introduced a scheme for evaluating dense matching methods with the help of calibration data and ground truth. This scheme evaluates the resulting point cloud of the reconstruction by separating inliers from outliers using Least Median Squares outlier detection, and compares their numbers and ratios.

Thesis 2 Dense Affine Matching

I have designed, implemented and tested a novel region growing based dense matching method that compensates the distortion typical for wide baseline images.

2.1 I have created a novel template matching method (Affine Matching, AM) that extends search for affine parameters, and thus is usable in the presence of image distortion. The affine compensation reduces matching errors and in the same time determines the best affine transformation that approximates the distortion. The matching is ideal for wide baseline image matching, where the distortion between the views is high.

2.2 Observing the properties of the approximating affine transformation on different surfaces I formulated novel matching

constraints (affine constraints) that can be used for dense matching.

2.3 I introduced a novel region growing based dense matching solution (Dense Affine Matching, DAM) that extends classical matching functions and uses affine constraints. Both propagation and termination rules in the region growing are affected by the affine constraints. The resulting 3D surfaces are smoother, denser, and more reliable on wide baseline input.

2.4 I evaluated the Dense Affine Matching with both real world images and quantitative evaluation methods. DAM proved superior to the method of Hirschmuller et al. in terms of precision, density and robustness.

Thesis 3 Normal Aided Matching

I have introduced a novel way to use calibration data in dense matching. By exploiting the connection between image distortion and surface orientation, I developed an improved region growing based dense matching method that uses surface normals.

3.1 I have formalized the connection between the affine distortion on different image views and the surface normals. The surface normal for a 3D patch can be calculated by measuring the distortion between the views. The method is applicable if linear rectification transformation is used and certain camera parameters are known.

3.2 Using the surface normals I formulated novel constraints that can be used in dense matching algorithms to reduce the search

space. The new constraints are connected to the physical geometry of surfaces and are unbiased with respect to surface orientation and viewing position.

3.3 I developed a novel region growing based dense matching method (Normal Aided Matching, NAM) that uses the new surface normal related constraints for termination and propagation rules. Whit this matching method, we found a new way to exploit calibration data during dense matching.

3.4 I tested and evaluated the NAM method on real world images and using a quantitative evaluation method. The method proved superior to both Hirschmuller et al. and DAM matching methods in terms of precision and robustness.

Author's publications on the topic of the dissertation

[1] **Zoltán Megyesi**. Sztereó látásban használt képillesztési algoritmusok. In *Proc. Third Hungarian Conference on Image Processing and Pattern Recognition (NJSZT-KÉPAF)*, pages 21–30, 2002.

[2] D. Chetverikov, **Z. Megyesi**, Z. S. Janko, and J. Matas. Using periodic texture as a tool for wide-baseline stereo. In *Proc. 26th Workshop of the Austrian Association for Pattern Recognition*, pages 37–44, 2002.

[3] **Z. Megyesi** and D. Chetverikov. Affine dense matching for wide baseline stereo. In *Proc. Grafika 2003*, pages 109–114, 2003.

[4] N. Novák and **Z. Megyesi**. Building 3D models using calibrated stereo and structured light. In *Proc. Fourth Hungarian Conference on Image Processing and Pattern Recognition (NJSZT-KÉPAF)*, pages 213–220, 2004.

[5] Dmitry Chetverikov, **Zoltan Megyesi**, and Zsolt Janko. Finding region correspondences for wide baseline stereo. In *ICPR '04: Proceedings of the Pattern Recognition, 17th International Conference on (ICPR'04) Volume 4*, pages 276–279, Washington, DC, USA, 2004. IEEE Computer Society.

[6] **Zoltan Megyesi** and Dmitry Chetverikov. Affine propagation for surface reconstruction in wide baseline stereo. In *ICPR '04: Proceedings of the Pattern Recognition, 17th International Conference on (ICPR'04) Volume 4*, pages 76–79, Washington, DC, USA, 2004. IEEE Computer Society.

[7] **Z. Megyesi** and D. Chetverikov. Enhanced surface reconstruction from wide baseline images. In *3D Data Processing, Visualization and Transmission, 2004. 3DPVT 2004. Proceedings. 2nd International Symposium on*, pages 463–469, 2004.

[8] **Z. Megyesi**, G. Kós, and D. Chetverikov. Surface normal aided dense reconstruction from images. In *Proc. Computer Vision Winter Workshop*, pages 64–69, 2006.

[9] **Zoltán Megyesi**, Géza Kós, and Dmitry Chetverikov. Dense 3D reconstruction from images by normal aided matching. *Machine GRAPHICS & VISION*, 15(1):3–28, 2006.

[10] **Zoltán Megyesi**. Kép alapú 3D modellalkotás forgóasztalhoz kapcsolt kamera rendszerrel. In *AGTEDU 2008*, volume 1, pages 248–253, 2008.

[11] **Zoltán Megyesi** and Gábor Kátai-Urbán. Kép alapú 3D modellalkotás forgóasztal és felületi normálissal javított illesztés használatával. In *Proc. Seventh Hungarian Conference on Image Processing and Pattern Recognition (NJSZT-KÉPAF)*, 2009.

[12] P. Kovács, Z. Gaal, A. Barsi and **Z. Megyesi**. Real time natural 3D contentdisplaying with holovizio displays. In *29th International Display Research Conference, EuroDisplay 2009*, September 2009. to appear.

[13] **Zoltán Megyesi**. Kép alapú 3D felület rekonstrukció széles bázistávolságú képek esetén, felületi normális meghatározásával. *GÉP, A Gépipari Tudományos Egyesület Országos Műszaki Folyóirata*, 2009. to appear.

Author's other publications

[**14**] D. Chetverikov, **Z. Megyesi**, and Z. S. Janko. An overview of research activities of image and pattern analysis group of mta sztaki. In *Proc. International Workshop East-West Vision 2002*, pages 35–40, 2002.

[**15**] T. Balogh, P. Kovács, and **Z. Megyesi**. Holovizio 3D display system. In *Proc. First International Conference on Immersive Telecommunications (IMMERSCOM 2007)*, 2007.

[**16**] Tomas Rodriguez, Adolfo C. de Leon, Bruno Uzzan, Nicolas Livet, Edmond Boyer, Florian Geffray, Tibor Balogh, **Zoltan Megyesi**, and Attila Barsi. Holographic and action capture techniques. In *SIGGRAPH '07: ACM SIGGRAPH 2007 emerging technologies*, New York, NY, USA, 2007. ACM.

[**17**] Tibor Balogh, Peter T. Kovacs, **Zoltan Megyesi**, and Attila Barsi. Holovizio true 3D display system. In *NEM Summit 2008*, 2008.

[**18**] **Z. Megyesi**, A. Barsi, and T. Balogh. 3D video visualization on the holovizio system. In *3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video, 2008*, pages 269–272, 2008.